



0590  
1023

**RAW SEQUENCE LISTING**  
**ERROR REPORT**  
**BEST AVAILABLE COPY**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/888,182

Source: O/PE

Date Processed by STIC: 11/1/2001

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

**FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.**

**FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.**

**PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:**

**Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

**Checker Version 3.0 can be down loaded from the USPTO website at the following address:**

**<http://www.uspto.gov/web/offices/pac/checker>**

OIEP

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/888,182

DATE: 11/01/2001  
TIME: 14:15:17

Input Set : A:\ES.txt  
Output Set: N:\CRF3\11012001\I888182.raw

Does Not Comply  
Corrected Diskette Needed

## ERRORED SEQUENCES

W--> 2 <210> SEQ ID NO: 1  
3 <211> LENGTH: 27  
4 <212> TYPE: DNA  
5 <213> ORGANISM: Artificial Sequence  
W--> 6 <220> FEATURE:  
7 <223> OTHER INFORMATION: Description of Artificial Sequence: Erk5-specific  
8 primer  
W--> 0 <110> APPLICANT:  
W--> 0 <120> TITLE INVENTION:  
W--> 0 <160> NUMBER OF SEQ ID NOS:  
W--> 9 <400> SEQUENCE: 1  
E--> 10 cagccattcg atgtgggccc acgcta

26

see following pages

Please consult Sequence Rules  
for valid format.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/888,182

DATE: 11/01/2001

TIME: 14:15:18

Input Set : A:\ES.txt

Output Set: N:\CRF3\11012001\I888182.raw

L:2 M:283 W: Missing Blank Line separator, <210> field identifier  
 L:6 M:283 W: Missing Blank Line separator, <220> field identifier  
 L:0 M:282 W: Numeric Field Identifier Missing, <110> is required.  
 L:0 M:282 W: Numeric Field Identifier Missing, <120> is required.  
 L:0 M:282 W: Numeric Field Identifier Missing, <160> is required.  
 L:9 M:283 W: Missing Blank Line separator, <400> field identifier  
 L:10 M:252 E: No. of Seq. differs, <211>LENGTH:Input:27 Found:26 SEQ:1  
 L:16 M:283 W: Missing Blank Line separator, <220> field identifier  
 L:19 M:283 W: Missing Blank Line separator, <400> field identifier  
 L:26 M:283 W: Missing Blank Line separator, <220> field identifier  
 L:30 M:283 W: Missing Blank Line separator, <400> field identifier  
 L:31 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
 L:37 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
 L:41 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
 L:42 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
 L:43 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
 L:44 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
 L:50 M:283 W: Missing Blank Line separator, <220> field identifier  
 L:54 M:283 W: Missing Blank Line separator, <220> field identifier  
 L:58 M:283 W: Missing Blank Line separator, <400> field identifier  
 L:59 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4  
 L:60 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4  
 L:61 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4  
 L:62 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4  
 L:65 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4  
 L:78 M:283 W: Missing Blank Line separator, <220> field identifier  
 L:82 M:283 W: Missing Blank Line separator, <400> field identifier  
 L:83 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
 L:85 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
 L:88 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
 L:99 M:283 W: Missing Blank Line separator, <220> field identifier  
 L:106 M:283 W: Missing Blank Line separator, <400> field identifier  
 L:107 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
 L:108 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
 L:109 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
 L:110 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
 L:111 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
 L:112 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
 L:113 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
 L:0 M:203 E: No. of Seq. differs, <160> Number Of Sequences:Input (0) Counted (6)

L1107  
L1207  
L1407  
L1417  
L1607

These mandatory numeric identifiers and responses are missing from the beginning of the Sequence Listing

~~SEQUENCE LISTING~~

<210> 1

<211> (27) 26 shown

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Erk5-specific primer

<400> 1

cagccattcg atgtggggccc acgcta

26

See next page for more errors

<210> 3  
<211> 802  
<212> DNA  
<213> Mus musculus  
<220>  
<221> misc\_feature  
<222> (1)..(38)  
<223> partial sequence of exon 2 from Erk5 gene  
<400> 3

cggnacctac tgtgccctat ggaggaattc agatctgtgt aagggagtgg gccaggagga 60  
ggagacacag tcgggatcag cttagaagcc caggttcagt aatactgaag ttctggcagg 120  
gcggttgaac ccagagtgat gcgggctgtg agtccaggac attggtaggg acagttctta 180

(partial listing of seq. 3)

ps1 Use of n and/or Xaa has been detected in the Sequence Listing.  
Review the Sequence Listing to insure a corresponding  
explanation is presented in the <220> to <223> fields of  
each sequence using n or Xaa.

→ see  
item 9  
on Enol  
summary  
sheet  
more n's  
are shown  
in seq. 3

see sample Sequence Listing (attached)

<110> Smith, John; Smithgene Inc.

<120> Example of a Sequence Listing

<130> 01-00001

<140> PCT/EP98/00001

<141> 1998-12-31

<150> US 08/999,999

<151> 1997-10-15

<160> 4

<170> PatentIn version 2.0

<210> 1

<211> 389

<212> DNA

<213> Paramecium sp.

<220>

<221> CDS

<222> (279)...(389)

<300>

<301> Doc. Richard

<302> Isolation and Characterization of a Gene Encoding a  
Protease from Paramecium sp.

<303> Journal of Genes

<304> 1

<305> 4

<306> 1-7

<307> 1988-06-31

<308> 123456

<309> 1988-06-31

<400> 1  
agctgtagtc attcctgtgt cctcttctct ctgggcttct caccctgcta atcagatctc 60  
agggagagtg tcttgacctt cctctgcctt tgcagcttca caggcaggca ggcaggcagc 120  
cgatgtggca attgctggca gtgccacagg cttttcagcc aggccttaggg tgggttcgcg 180  
cgcgggcgcgg cggccctctt cgcgcctctc tcgcgcctct ctctcgtctt cctctcgtct 240

Appendix 3, page 2

ggacctgatt	aggtgagcag	gaggaggggg	cagtttagc	atg	ggt	tca	atg	ttc	agc	296
				Met	Val	Ser	Met	Phe	Ser	
				1				5		
ttg	tct	ttc	aaa	tgg	cct	gga	ttt	tgt	ttg	344
Leu	Ser	Phe	Lys	Trp	Pro	Gly	Phe	Cys	Leu	
			10					15	20	
tgt	ccc	aaa	gtc	ctc	ccc	tgt	cac	tca	tca	
Cys	Pro	Lys	Val	Leu	Pro	Cys	His	Ser	Ser	
		25					30			
<210>		2								
<211>		37								
<212>		PRT								
<213>		Paramecium sp.								
<400>		2								
Met	Val	Ser	Met	Phe	Ser	Leu	Ser	Phe	Lys	
1				5					10	
Phe	Val	Cys	Leu	Phe	Gln	Cys	Pro	Lys	Val	
			20					25		
Leu	Gln	Pro	Asn	Leu						
			35							
<210>		3								
<211>		11								
<212>		PRT								
<213>		Artificial Sequence								
<220>										
<223>										
<400>		3								
Met	Val	Asn	Leu	Glu	Pro	Met	His	Thr	Glu	Ile
1				5					10	
<210>		4								
<400>		4								
000										

[Annex VIII follows]

identifiers and their accompanying information as shown in the following table. The numeric identifier shall be used only in the "Sequence Listing." The order and presentation of the items of information in the "Sequence Listing" shall conform to the arrangement given below. Each item of information shall begin on a new line and shall begin with the numeric identifier enclosed in angle brackets as shown. The submission of those items of information designated with an "M" is mandatory. The submission of those items of information designated with an "O" is optional. Numeric identifiers <110> through <170> shall only be set forth at the beginning of the "Sequence Listing." The following table illustrates the numeric identifiers.

Numeric Identifier	Definition	Comments and Format	Mandatory (M) or Optional (O)
<110>	Applicant	Preferably max. of 10 names; one name per line; preferable format: Surname, Other Names and/or Initials	M
<120>	Title of Invention		M
<130>	File Reference	Personal file reference	M, when filed prior to assignment of appl. number
<140>	Current Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if available
<141>	Current Filing Date	Specify as: yyyy-mm-dd	M, if available
<150>	Prior Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if applicable include priority documents under 35 USC 119 and 120
<151>	Prior Application Filing Date	Specify as: yyyy-mm-dd	M, if applicable
<160>	Number of SEQ ID NOs	Count includes total number of SEQ ID NOs	M
<170>	Software	Name of software used to create the Sequence Listing	O
<210>	SEQ ID NO: #:	Response shall be an integer representing the SEQ ID NO shown	M
<211>	Length	Respond with an integer expressing the number of bases or amino acid residues	M



<212>	Type	Whether presented sequence molecule is DNA, RNA, or PRT (protein). If a nucleotide sequence contains both DNA and RNA fragments, the type shall be "DNA." In addition, the combined DNA/RNA molecule shall be further described in the <220> to <223> feature section.	M
<213>	Organism	Scientific name, i.e. Genus/species, Unknown or Artificial Sequence. In addition, the "Unknown" or "Artificial Sequence" organisms shall be further described in the <220> to <223> feature section.	M
<220>	Feature	Leave blank after <220>. <221-223> provide for a description of points of biological significance in the sequence.	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.
<221>	Name/Key	Provide appropriate identifier for feature, preferably from WIPO Standard ST.25 (1998), Appendix 2, Tables 5 and 6	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence
<222>	Location	Specify location within sequence; where appropriate state number of first and last bases/amino acids	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified

		in feature	base was used in a sequence
<223>	Other Information	Other relevant information; four lines maximum	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.
<300>	Publication Information	Leave blank after <300>	0
<301>	Authors	Preferably max of ten named authors of publication; specify one name per line; preferable format: Surname, Other Names and/or Initials	0
<302>	Title		0
<303>	Journal		0
<304>	Volume		0
<305>	Issue		0
<306>	Pages		0
<307>	Date	Journal date on which data published; specify as yyyy-mm-dd, MMM-yyyy or Season-yyyy	0
<308>	Database Accession Number	Accession number assigned by database including database name	0
<309>	Database Entry Date	Date of entry in database; specify as yyyy-mm-dd or MMM-yyyy	0
<310>	Patent Document Number	Document number; for patent-type citations only. Specify as, for example, US 07/999,999	0

<311>	Patent Filing Date	Document filing date, for patent-type citations only; specify as yyyy-mm-dd	0
<312>	Publication Date	Document publication date, for patent-type citations only; specify as yyyy-mm-dd	0
<313>	Relevant Residues	FROM (position) TO (position)	0
<400>	Sequence	SEQ ID NO should follow the numeric identifier and should appear on the line preceding the actual sequence	0

5. Section 1.024 is revised to read as follows:

1.024 Form and format for nucleotide and/or amino acid sequence submissions in computer readable form.

(a) The computer readable form required by 1.021(c) shall meet the following specifications:

(1) The computer readable form shall contain a single "Sequence Listing" as either a diskette, series of diskettes, or other permissible media outlined in paragraph (c) of this section.

(2) The "Sequence Listing" in paragraph (a) (1) of this section shall be submitted in American Standard Code for Information Interchange (ASCII) text. No other formats shall be allowed.

(3) The computer readable form may be created by any means, such as word processors, nucleotide/amino acid sequence editors or other custom computer programs; however, it shall conform to all specifications detailed in this section.

(4) File compression is acceptable when using diskette media, so long as the compressed file is in a self-extracting format that will decompress on one of the systems described in paragraph (b) of this section.

(5) Page numbering shall not appear within the computer readable form version of the "Sequence Listing" file.

(6) All computer readable forms shall have a label permanently affixed thereto on which has been hand-printed or typed: the name of the applicant, the title of the invention, the date on which the data were recorded on the computer readable form, the operating system used, a reference number, and an application serial number and filing date, if known.

(b) Computer readable form submissions must meet these format requirements:

(1) Computer: IBM PC/XT/AT, or compatibles, or Apple Macintosh;

(2) Operating System: MS-DOS, Unix or Macintosh;